An Algorithm To Calculate The Facet Dimension Of A Bell Inequality

Author: Neil Smith

Supervisors: Mark Howard and Pieter Kok

August 2017

Abstract

The study of quantum theory has led to some of the most innovative technologies to date and current research on the topic of non-locality is leading to promising applications in information processing and communication complexity tasks. In order to understand how useful this feature could be, it is important to address the problem of characterising the local set of correlations of a general scenario. To date, this has only been done for a few simple systems, however a good characterisation can be provided by non-trivial Bell Inequalities found by the calculation of robustness. In this work we present an algorithm that calculates the facet dimension of a general Bell Inequality which is vital in evaluating the usefulness of these Inequalities.

Contents

1 Introduction

To date, quantum mechanics has been used in many applications, exploiting its peculiar properties, enabling us to achieve things that would have otherwise been impossible or impractical to do with traditional classical theory and technologies. Over the last several decades one particular feature of quantum mechanics, non-locality, has become a fruitful topic of research. It has the potential to enable the accomplishment of new communication and information-processing tasks and render traditionally difficult tasks easy.

The concept of non-locality dates back to 1935 where Albert Einstein, Boris Podolsky and Nathan Rosen (EPR) presented a paper [9] in which they argued that quantum mechanics is incomplete, and proposed that it could be completed through the use of hidden variables, a local hidden variable theory. However, later in 1964, John Bell constructed an expression, a Bell Inequality, that could be used to experimentally test hidden variable theories through the use of measurable physical quantities [2]. Shortly thereafter, Bell proved that no local hidden variable theory could reproduce all of the predictions of quantum mechanics, this became known as Bell's theorem [3]. Since then, there have been many experiments that have been used to show that nature does indeed not behave in the manner described by these classical hidden variable theories. [1]

In their argument, EPR incorrectly assumed that signals cannot be transmitted between parties at speeds greater than the speed of light, this is known as the assumption of locality. This seems like a reasonable assumption to make, as it would be in contradiction with the theory of relativity. However, it has be observed that in some situations systems can interact in a way such that, when they are then separated, they can transmit signals between themselves faster than the speed of light, in fact this transmission is instantaneous. This phenomenon is known as non-locality and arises due to the entanglement of the systems. This appears like it would enable us to transmit information faster than the speed of light, but in reality this phenomenon cannot be exploited in this way as the signals that are sent are in fact random signals. Correlations can be observed between distant parties, but these correlations are correlations of randomness and so no information can be transmitted.

However, non-locality does still have other useful applications, particularly in communication complexity tasks. The entanglement of a quantum system can be exploited to perform tasks that are expensive or even impossible to do with classical systems. In order to understand how useful this phenomenon could be, it is important to be able to calculate the maximum classical and quantum correlations observable in a particular scenario. This project deals with the first of these problems. In order to calculate the maximum classical correlation one needs to derive the tight Bell Inequalities that characterise the local set of correlations, this has only been done for a few very simple systems. However, methods have been developed to derive non-trivial Bell Inequalities, from the calculation of robustness [11], which could still be useful for characterising the local set. The task here then, is to measure how good at characterising the local set these Inequalities are. One way of doing this is to check whether or not the Inequality corresponds to a facet by calculating how many of the vertices of the polytope lie on the Inequality and are linearly independent. If there are no vertices that lie on the Bell Inequality then we know that it does not correspond to a facet, if there are, then this number tells us the dimension of this facet Bell Inequality. Currently, there is no general algorithm to calculate whether a given Bell Inequality is a facet, here we present such an algorithm. In section 2 the reader will be taken through the necessary material to understand the topic and in section 3 the reader will be presented with the algorithm with a short discussion on its functions and implementation. In section 4 we discuss how well the algorithm works and the different tests it passed and in section 5 we summarise our findings.

2 Preliminaries

2.1 Nonlocality

Here the reader will be introduced to a framework in which we can study nonlocality. Consider the following, there are n parties who are very far apart[1] from each other, and each have access to their own system upon which they can make measurements and observe an outcome. In general, the number of possible measurements that each party can make upon their system will vary and so will the number of outcomes for each of their measurements. In this work, however, we will only concern ourselves with the special case in which all of the parties have the same number of measurements they can perform m and each party has the same number of outcomes they can observe for each of their measurements d. This is known as a scenario, each scenario is therefore specified by the three numbers (n, m, d). In this framework we only care about the number of possible measurements and outcomes that could occur, not the measurements and outcomes themselves or the inner workings of the systems, this is known as a device-independent scenario; each system can be treated as a black box with just a series of possible inputs and outputs. Taking this further, party i can make a measurement m_i from a set of possible measurements M_i and observe a corresponding outcome d_i from a set of possible outcomes D_i.[2] Figure 1 shows such a scenario with just two parties, Alice and Bob.

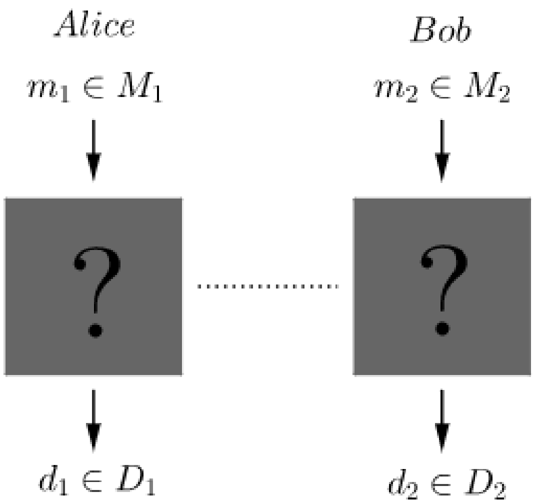

Figure 1: The Alice and Bob scenario, there are two parties who are very far apart from each other and each have access to a system upon which they can make measurements and observe outcomes. Alice can make a measurement m_1 on her system from a set of possible measurements M_1 and observe an outcome d_1 from a set of possible outcomes D_1. Similarly, Bob can make a measurement m_2 on his system from a set of possible measurements M_2 and observe an outcome d_2 from a set of possible outcomes D_2.

In each scenario we can define a joint probability $P(d_1d_2...d_n|m_1m_2...m_n)$ that the parties make these measurements and observe these outcomes.[3] We can also define a corresponding joint probability distribution, in the case that there are a finite number of measurements and outcomes that could occur, this joint probability distribution can be represented as a vector (1) with $(md)^n$ components, which is known as a behaviour as it represents a possible behaviour of the scenario.

$$\vec{P}(d_1d_2...d_n|m_1m_2...m_n) \in \mathbb{R}^{(md)^n} \tag{1}$$

In the field of nonlocality we are interested in what behaviours, and what correlations between the

[1] Technically, space-like separated.

[2] Since we do not concern ourselves with the measurements or outcomes, we can just assume that the set of possible outcomes is the same for each measurement of party i and denote this set D_i.

[3] It is important to note that we also assume they can only make one measurement at a time, the reason for this becomes clear later on when we study quantum behaviours, certain measurements do not commute with each other.

measurements could occur, given different conditions on how the parties can communicate and what systems they have access to.

2.1.1 Local Correlations

Given that the parties are far apart and cannot communicate, you might expect that the only way for there to be a non-zero correlation between each party's measurements is if there exists some common cause λ in the past that pre-determines the outcomes that could occur. This is the view taken in a local hidden variable theory, the cause is the hidden variable $\lambda \in \Lambda$ that gives rise to the correlations they observe. If we assume that this is the reason that the results are correlated then there are only a certain set of behaviours and correlations that can be observed. These behaviours are known as the local behaviours, all local behaviours can be expressed in the form (2), where $q(\lambda) \geq 0$ is a probability distribution and $\int q(\lambda)\, d\lambda = 1$.

$$P(d_1 d_2 ... d_n | m_1 m_2 ... m_n) = \int_\Lambda q(\lambda) p(d_1|m_1\lambda) p(d_2|m_2\lambda) ... p(d_n|m_n\lambda)\, d\lambda \tag{2}$$

The local correlations $\mathcal{L}$ are a subset of the local behaviours that satisfy the normalization (3) and no-signalling conditions (8) and therefore represent real physical behaviours of systems. The no-signalling conditions will be discussed in detail later in section 2.1.3. From now on when we refer to the "local behaviours" we are really referring to these local correlations. Any correlation that cannot be explained through a local hidden variable theory is known as a non-local correlation.

$$\sum_{d_1, d_2, ..., d_n} P(d_1 d_2 ... d_n | m_1 m_2 ... m_n) = 1 \quad \forall \quad m_i \in M_i \quad \forall \quad i \in \{1, ..., n\} \tag{3}$$

There are three main assumptions made about the scenario in a local hidden variable theory, these are:

1. Locality: The systems are influenced only by their local surroundings.
2. Realism: The properties of systems have well-defined values independent of external measurements.
3. Free Will: The parties can freely choose which measurements to perform.

The first, also known as the assumption of local causality, is the important one here. We assume that the parties are far enough apart that they cannot influence or communicate with each other, by sending a signal. This signal must therefore have a finite speed. In order for the local hidden variable theory to be consistent with the special theory of relativity, the speed of this signal cannot be greater than the speed of light, in order words the systems are unaffected by anything outside their light cones. Another way to phrase this is that each party's choice does not depend on the measurement choices of the other parties and the choice of which measurement to make is made when they are far apart. The local probability distribution of one party's outcome is independent of the experiments performed by the other parties. Mathematically, this can be expressed for party 1 as (4).

$$P(d_1 | d_2 ... d_n m_1 m_2 ... m_n \lambda) = P(d_1 | m_1 \lambda) \quad \forall\, \lambda \in \Lambda \tag{4}$$

You may ask what is the purpose of the second assumption, isn't this always true? This is what we are used to thinking in classical physics, but in quantum mechanics the properties of systems do not always have well-defined values until they are measured. In fact, the result depends on the specific experimental setup being used to measure it, this is known as quantum contextuality. In the assumption of free will we assume that the common cause λ is not correlated with their choice of measurements. In other words their measurement choices do not depend on the cause of the correlations, either directly or indirectly. This is also known as the assumption of measurement independence.

2.1.2 The Local Polytope

Earlier, it was mentioned that the local behaviours are defined according to (2) but they can be equivalently defined through (5), where $c_i > 0$ and $\sum_i c_i = 1$. The probabilities can be expressed as a convex sum of products of deterministic probability[4] distributions of each party. Therefore, the local set of behaviours forms a convex set and in the scenarios where the parties can only make a finite number of measurements and observe a finite number of outcomes, this set forms a convex polytope, known as the local polytope. A polytope is a convex set with a finite number of extremal points[5], see Figure 2.

$$\mathcal{L} = \left\{ \vec{P}(d_1...d_n|m_1...m_n) \,|\, P(d_1...d_n|m_1...m_n) = \sum_i c_i D_1(d_1|m_1) D_2(d_2|m_2)...D_n(d_n|m_n) \right\} \tag{5}$$

The extremal points of the polytope are the behaviours part of the local set that can be written as a single product of the deterministic probability distributions of each party (6), because the set is a convex polytope these can be used to characterise the set. In total, for a scenario (n, m, d) there will be d^{nm} of these extremal behaviours. [13]

$$P(d_1 d_2...d_n|m_1 m_2...m_n) = D(d_1|m_1) D(d_2|m_2)...D(d_n|m_n) \tag{6}$$

The dimension of the local polytope (7) can be found by considering the normalization and no-signalling conditions. [14]

$$(m\,(d-1)+1)^n - 1 \tag{7}$$

2.1.3 No-Signalling Correlations

Local hidden variable theories impose many more assumptions upon the scenario than just the assumption that the parties cannot communicate faster than the speed of light. We can also consider what behaviours could occur given just this assumption about the parties, this is known as a no-signalling scenario. We assume that the parties are far enough apart that they cannot get any information about the other parties' experiments, this means that the probability of one party getting a given outcome is independent of any of the other parties' experiments. Mathematically this can be expressed as (8), which is equivalent to requiring that the marginals of the joint probability distribution are well-defined.

$$\sum_{d_k} P(d_1...d_k...d_n|m_1...m_k...m_n) = \sum_{d_k} P(d_1...d_k...d_n|m_1...m_k'...m_n) \tag{8}$$
$$\forall\ m_k, m_k' \in M_K, \quad \forall\ k \in \{1, ..., n\}, \quad \forall\ d_i, m_i \in D_i, M_i\ (i \neq k)$$

The behaviours that satisfy these no-signalling conditions form the no-signalling set, this also forms a convex polytope. Since these conditions are less restrictive than the conditions for the local set, the local set of behaviours is contained within the no-signalling set (9), this inclusion is in fact strict. [15] We expect to see no-signalling correlations greater than that allowed by the local set.

$$\mathcal{L} \subset \mathcal{NS} \tag{9}$$

[4]A deterministic probability is one in which the probability of an outcome is either one or zero $D(d|m) \in \{0, 1\}$.

[5]The local polytope is also known as the convex hull of the local behaviours.

2.1.4 Quantum Correlations

A quantum behaviour is one that can be written in quantum theory as a trace (10) of a quantum state $\rho \in \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes ... \otimes \mathcal{H}_n$ where $\{M_{m_1}^{d_1}, M_{m_2}^{d_2}, ..., M_{m_n}^{d_n}\}$ are local POVMs acting on the arbitrary Hilbert spaces $\{\mathcal{H}_1, \mathcal{H}_2, ..., \mathcal{H}_n\}$ respectively. The set of all quantum behaviours is denoted $\mathcal{Q}$.

$$P(d_1 d_2 ... d_n | m_1 m_2 ... m_n) = \mathrm{Tr}\left(M_{m_1}^{d_1} \otimes M_{m_2}^{d_2} \otimes ... \otimes M_{m_n}^{d_n}\, \rho\right) \tag{10}$$

The quantum set is more restrictive than the no-signalling set but less restrictive than the local set, for this reason, the local set is contained within the quantum set[6] and the quantum set is contained within the no-signalling set (11), these inclusions are both strict. [15, 6] We can show that the quantum set is contained within the no-signalling set by showing they satisfy (8), this is done in (A.1) in the appendix.

$$\mathcal{L} \subset \mathcal{Q} \subset \mathcal{NS} \tag{11}$$

The quantum set is also a convex set like both the local and no-signalling sets but does not form a polytope, Figure 2 shows this relationship. Quantum behaviours can give rise to correlations greater than that allowed by a local hidden variable theory but less than that allowed by the no-signalling set. States ρ that are separable can only generate local correlations, whereas states that are entangled can generate correlations that cannot be observed in a local hidden variable theory.

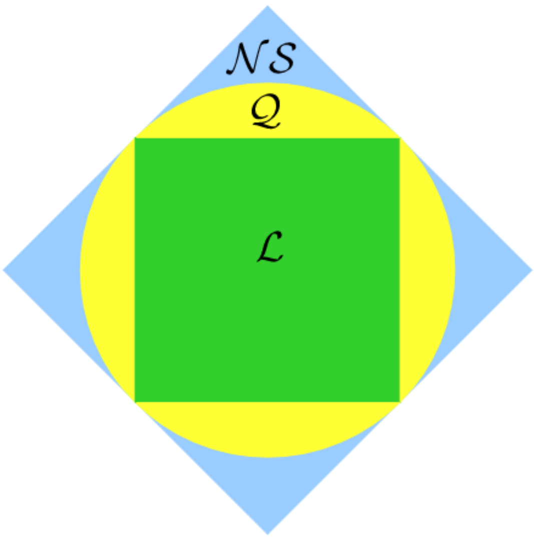

Figure 2: A schematic depicting the local $\mathcal{L}$, quantum $\mathcal{Q}$ and no-signalling $\mathcal{NS}$ sets of correlations. The local set is contained within the quantum set, which is contained within the no-signalling set, they are all convex sets. Both the local and no-signalling sets are polytopes, but the quantum set is not.

Currently, there is no known method to calculate the quantum bound, at least in a finite amount of time, in fact there are no set of principles that characterise the quantum set, and it is not known whether such principles exist. This is one of the main unsolved problems in the field on nonlocality. There have been many attempts to characterise the quantum set, one of the best attempts was achieved in 2008 when Navascués, Pironio and Acín found an infinite heirarchy of well characterised sets that converges to the quantum set, known as the NPA heirarchy. [12] The NPA hierarchy has been used to efficiently compute better and better upper bounds on the quantum bound. Methods have also been developed to calculate a lower bound, this is usually attained through the see-saw iterative method. [16]

[6]This is to be expected since classical physics is contained within quantum physics.

2.2 Bell Inequalities

We know that the extremal points of the local polytope can be used to characterize the local set, but the set can also be characterised through its facets or the intersection of the corresponding series of halfspaces. Bell Inequalities are expressions that can be used to characterise the local correlations that can be observed, they represent halfspaces. They take the form (12) where S is a linear sum of correlators and S represents a bound on the local correlations that could occur.[7]

$$S = \sum_i c_i E_i \leq S_{max} \tag{12}$$

The first Bell Inequality was introduced by Bell in his 1964 paper [2] as a way of providing a test of whether or not quantum mechanics could be completed through the use of hidden variables. A violation of a Bell Inequality provides evidence that at least one of the assumptions made in a local hidden variable theory is incorrect, indeed violations were observed and they showed that no completion in this way was possible. To date, experiments have ruled out realism, showing that nature behaves nonlocally.[8]

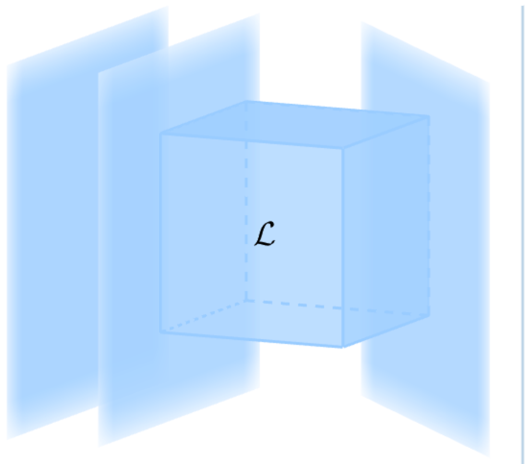

Figure 3: The geometric representation of Bell Inequalities as a series of hyperplanes or halfspaces in correlation space. Tight Bell Inequalities represent the facets of the local set of correlations. There are an infinite number of Bell Inequalities for each scenario, some represent hyperplanes parallel to the facet and others are not, some are closer to the facet and others are further away. Bell Inequalities can also have a lower dimension than that of the facet, these are represented as lines and vertices.

The parties each have a selection of measurements they could perform on their systems but they will not always perform each of their measurements and so the coefficients in (12) may not all be non-zero. Therefore, there are many different Bell Inequalities I that are possible within each scenario, in fact, there are an infinite number of Bell Inequalities, each one represents a different halfspace. Each Bell Inequality can also be defined through its hyperplane that defines the bound of the Inequality. A Bell Inequality that describes a facet of the local space of correlations is known as a tight Bell Inequality I_{facet}, the set of all these provide a minimal characterisation of the local set and are the most useful. However, not all of the non-trivial Bell Inequalities are completely useless, they can still be used to characterise the local set, to a certain degree. These non-trivial Bell Inequalities may be further away from the facet or they may be closer, they may be parallel to the facet or they may not be, they can even have a dimension lower [13] than that of the tight Bell Inequality (13), this is summarised in Figure 3. It is important to understand these properties in order to understand how useful a particular Bell Inequality is. The dimension of a tight Bell Inequality (13)

[7]Note that not all Bell Inequalities can be written as a sum of correlators, all Bell Inequalities can be written in terms of probabilities, but the ones considered here can.

[8]Free will has not been ruled out but it is in fact scientifically impossible to rule out free will.

will always be one less than that of the corresponding polytope (7).

$$\dim I \leq \dim I_{facet} = \dim \mathcal{L} - 1 = (m\,(d-1)+1)^n - 2 \tag{13}$$

Equation (14) shows the form of a correlator for a general scenario (n, m, d) for n parties where party i makes a measurement m_i on their system, this is expressed in terms of their local deterministic probabilities. As mentioned previously, terms can arise in Bell Inequalities where one or more of the parties do not make any measurement on their system, in this case the correlator in (14) takes a different form; the joint probability distribution is replaced with the n partite marginal over the parties that do not make measurements.[9]

$$E_{m_1 m_2 \ldots m_n} = \sum_{d_1, d_2, \ldots, d_n} (-1)^{d_1+d_2+\ldots+d_n-n} D_1(d_1|m_1) D_2(d_2|m_2) \ldots D_n(d_n|m_n) \tag{14}$$

In total there will be $(m+1)^n - 1$ possible correlators that could arise, these can be arranged into a vector $\vec{E}$. In order to specify a particular Bell Inequality one only needs to specify the coefficients of these correlators, these can then also be arranged into a vector of the same length $\vec{n}$ which represents the normal vector of the set of hyperplanes associated with the Bell Inequality, see Figure 4. This leads to (15) which is an equivalent formulation of a Bell Inequality, which clearly shows the interpretation of a Bell Inequality as a half space.[10]

$$\vec{E} \cdot \vec{n} \leq S \tag{15}$$

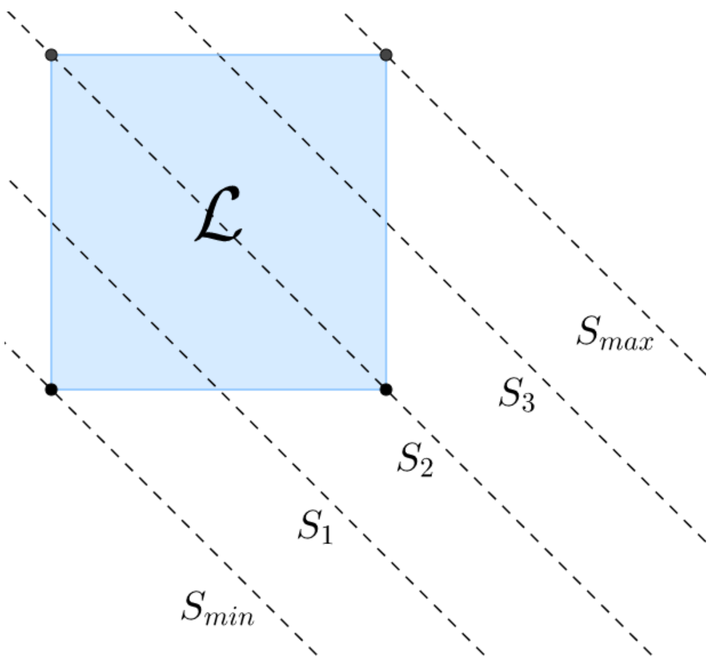

Figure 4: A depiction of a Bell Inequality as a series of parallel hyperplanes $\vec{E} \cdot \vec{n} = S$. The diagram shows the local correlations that can be obtained $S_{min} \leq S \leq S_{max}$ for this set of correlators $\vec{n}$. Each vector $\vec{n}$ specifies a different Bell Inequality by specifying the direction of the hyperplanes.

An example of a tight Bell Inequality is the CHSH Inequality [7], given in (16), it is one of the most famous

[9]Note that here the outcomes start at 1 instead of 0, this is to be consistent with the measurement notation being used here, party i making no measurement is denoted by $m_i = 0$. This also means everywhere the value of d_i appeared, we have to subtract one unit to be consistent, this is the origin of the term $-n$.

[10]It is important to note that it is more correct to say that the modulus of $\vec{E} \cdot \vec{n}$ is bounded by S this leads to two inequalities, but these are actually equivalent under a relabelling of the measurement outcomes of a single party. In fact, there are many equivalent Inequalities which can be found through a relabelling of the parties, measurements and outcomes.

and well-studied Bell Inequalities. It is the simplest tight Bell Inequality for which non-local correlations can be observed. It represents the correlations that can be observed between two parties in the scenario $(2, 2, 2)$.

$$E_{11} + E_{12} + E_{21} - E_{22} \leq 2 \tag{16}$$

2.3 Communication Complexity

Bell Inequalities are not only important in the study of nature, they also have useful applications in information processing tasks, in particular that of communication complexity. Communication complexity is the study of how successful a group of parties are in calculating the answer $f(d_1, d_2, ..., d_n)$ to some problem given that each party i only has access to part d_i of the resources $\{d_1, ..., d_n\}$ needed to calculate the answer, and that the communication between the parties is restricted[11]. There are two main questions that can be answered here:

1. Given that the resources are distributed in a certain way, what is the minimal amount of communication necessary in order to guarantee calculation of the correct answer?

2. Given a certain restriction on the amount of communication, what is the probability of calculating the correct answer?

The latter of these is less intuitive, the concept of the probability of success does not exist in a classical computation, but does in a quantum computation.

The simplest and most obvious way to ensure calculation of the correct answer is to just send all of the information to one of the parties, however we are interested in finding methods which result in the answer with less communication. We would also like to know if communication complexity tasks exist where using an entangled system with quantum correlations would allow us to complete the task with less resources than if we had used a classical system, and by how much this would reduce the resources required.

It has been shown that every Bell Inequality is in fact linked to a corresponding communication complexity task [4]. The greater the amount of violation of the Bell Inequality by the quantum bound the larger the advantage of using the quantum system over a classical system. Therefore, there is a lot of ongoing research into finding systems where the difference between the classical and quantum bound is large, but, as mentioned earlier in section 2.1.4, we run into a problem as there is no known method of calculating the quantum maximum correlation, only methods for calculating lower and upper bounds on this value have been developed.

The most clear example of a communication complexity task is circuit design. Here each of the processors computes the answer to a small part of the overall calculation, we would like to know how we could minimise the communication between the processors in order to reduce the time and energy costs and creating a more efficient circuit design. Other examples of communication complexity tasks include the study of data structures and the optimization of networks. The study of Bell Inequalities also has important applications in cryptography and the security of communication. [10]

[11] We are only interested in the amount of communication necessary, not the number of computational steps involved nor the amount of memory required to store the information.

3 The Algorithm

As previously mentioned, an important application of non-locality is its use in communication complexity tasks, where we want to know the classical bound on the correlations of a particular tight Bell Inequality. However, in general, for a given scenario (n, m, d), it is difficult to derive the corresponding tight Bell Inequalities, in fact, tight Bell Inequalities have only been derived for the simplest of scenarios. However, a good characterisation of the local set of correlations can be provided by non-trivial Bell Inequalities. Methods have been developed to create non-trivial Bell Inequalities from states by calculating the robustness of these states, which is a linear programming problem[12], and so can be easily computed.

The purpose of this algorithm is to check whether or not these Inequalities can provide a good characterisation of the local set. It calculates how many of the vertices of the local polytope lie on the Bell Inequality. If none of the vertices lie on the Bell Inequality then it is not that useful, but if some do then we know that this Bell Inequality represents a facet of the local polytope, although it may have a dimension lower than the best possible dimension. If the number of vertices is non-zero then this number represents how close this Bell Inequality is to being a tight Bell Inequality. Currently, a general algorithm for checking such a property does not exist, the purpose of this work is to create such an algorithm.

As shown earlier in (6) the joint probability distribution at these extremal points can be written as a product of the single party local deterministic probability distributions. In general, if there is a scenario (n, m, d) with n parties who can all perform the same number of measurements m and each have the same number of outcomes d then there will be md components to each party's local deterministic probability distribution, as shown in (17). In total then, there will be nmd of these components which can be treated as variables. From now on, these single party local deterministic probabilities will be referred to as "the variables", each one is specified by the party n_i, the outcome d_i and the setting m_i.

$$\vec{D}(a|x) = (D(1|1), ..., D(1|m), D(2|1), ..., D(2|m), ..., D(d|m)) \tag{17}$$

Each of these variables can take either one of the values zero or one, this leads to a total of 2^{nmd} possible extremal behaviours, but actually not all of these are behaviours because they do not obey the normalisation (3) and no-signalling (8) conditions. Therefore, we have to check which ones obey the conditions, this task can be simplified by combining the conditions with (6) to end up with one simpler constraint to check (18). The derivation of this constraint is done within the appendix (A.2).

$$\sum_{d_k} D_k(d_k|m_k) = 1 \quad \forall\, m_k \in M_k \quad \forall\, k \in \{1, ..., n\} \tag{18}$$

The task reduces to calculating the expression (18) for all of these 2^{nmd} possible combinations of these variables. However, instead of doing this, we can make use of (18) to speed this task up. We know that there should be a total of d^{nm} extremal behaviours, and this can be shown by studying (18), in fact, we can know what these states are. If we look at (18), we can see that in general there are nm constraints which will each consist of a sum of variables. But since each constraint contains variables that do not appear in any of the other constraints and that each of these variables can only be either 0 or 1 and that the sum of each constraint must be unity, we can see that only one of the variables in each constraint can take the value 1. Since there are d variables in each constraint, there are only d ways in which this could happen. Since there are nm independent constraints, there will be d^{nm} ways in which you can arrange all the 1s amongst the variables such that they obey the constraints. Therefore, we know what these states are and we just have to search over the space of d^{nm} extremal behaviours instead of 2^{nmd}.

To implement this search, we can define another set of variables which we will call the index variables. There will be nm of these index variables and their values tell us which of the variables in the expressions (18) take the value 1. The value of the kth index variable tells us which one of the variables in the kth constraint is 1. Instead of looping over all the possible values of the variables we instead loop over the possible values

[12]The robustness is a measure of how far a particular state is from the local set. For a more detailed discussion of robustness see [11].

of these index variables and then find the values of the variables from these.

The algorithm should be able to take any Bell Inequality from any scenario and calculate the degree to which it is a facet. Here we will call this the facet dimension or more simply the dimension of the Bell Inequality, in the case that it is a tight Bell Inequality this will be the dimension of the Bell Inequality, these should not be confused. At its simplest one could just write out a series of nested for loops for each of the index variables, but in order to have generality the algorithm should be able to dynamically create an arbitrary number of for loops. This can be done through the use of recursion. The function takes as parameters the scenario n,m and d and a list of coefficients of correlators, the vector $\vec{n}$ in (15), which determines the particular form of the Bell Inequality.

A function is called with the number of index variables to loop over as a parameter. If there are variables that still have to be looped over then it will loop over one of these and within this for loop call the function again but with one less variable to loop over. If there are no more variables to loop over then it will calculate the current values of the variables and start the calculation of S. This process can be visualised in Figure 5 as a tree search. Since there are nm index variables and each one can take d different values, the tree will have a depth of nm with a branching factor d. This search is exponential in time but the only things that have to be stored in memory are the current values of the variables and the values of the variables that give the bound, this will always be a fraction of d^{nm}, the total number of vertices.

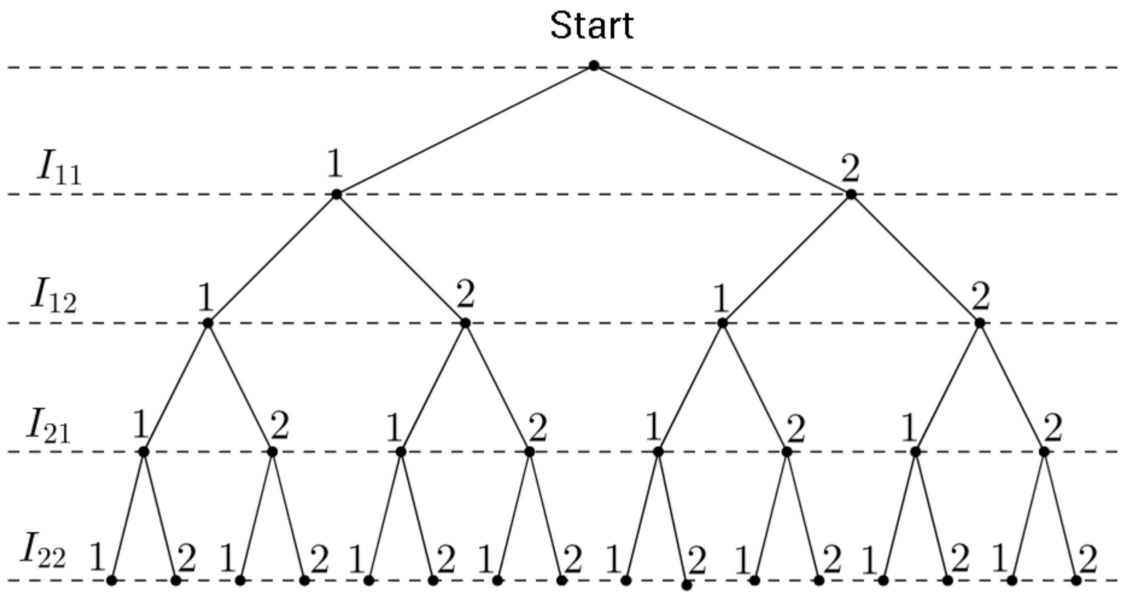

Figure 5: The index variable search space of the algorithm for the CHSH Inequality of scenario $(2, 2, 2)$. The search space takes the form a tree diagram. It has a branching factor of $d = 2$ and a depth of $nm = 4$. Each node represents a particular value of one of the nm index variables. Nodes at the same depth represent different values of the same index variable. Nodes at different depths represent different index variables.

However there is a problem with a simple recursion in that at any given time there will be nmd copies of the same list of variables and other quantities required to perform the calculation. This is a large cost to the memory. Instead, a better method would be to use an OOP approach, turning the function into a handle class. This would mean that these quantities can be made properties ensuring that there is only one copy of these quantities, and the values of these can be accessed and modified at any point in the algorithm. Making a handle class gives the nested methods access to the same object instead of copies of the object, this is necessary for the purposes of this program.

The coding of this project was done in Matlab and a list of the methods that were implemented and what they do are given in Table 1. The algorithm consists of just two files, the first contains the class cbanddimcalc.m, which is really just the algorithm itself. The second file calcdimandclassicalbound.m is just a function that is used to encapsulate this class. A more detailed description of the algorithm and each method and its implementation is given within the documentation.

Method	Description
Calc	Starts the calculation of the dimension and classical bound.
Loopdetprobs	Loops over the possible values of the variables.
Getdetprobindex	Calculates the index of the variable $D_{n_i}(d_i\|m_i)$ within the array of variables.
Calccorr	Calculates the correlator of a particular set of measurements $m_1 m_2 ... m_n$.
Calcprobdists	Calculates the joint probability distributions from the given sets of variables.
Calcprobdist	Calculates the joint probability distribution of a single set of variables.
Calcdim	Calculates the dimension from the joint probability distributions.

Table 1: The methods part of the class which calculates the classical bound and dimension of a Bell Inequality.

4 Testing And Results

The Algorithm was tested against a series of known tight Bell Inequalities for various scenarios and its time taken to calculate the facet dimension was measured. The program was tested against Inequalities with two, three and four parties, these are the CHSH inequality [7], equation (8) in [17], equation (19) in [8] and examples 2, 3 and 4 in [5]. As shown in Table 2 the program produced the correct results for the Inequalities with three or less parties but failed for the four party Inequalities. As a result, the program cannot be assumed to work correctly for Inequalities with $n \geq 4$. It is thought that the reason for this is that there is an error in the calculation of the bound as the program results in a value that is greater than the maximum for one of the Inequalities. However, for the other Inequality that it fails to produce the correct results for, it calculates the correct bound but with a dimension lower than expected. Nevertheless, an investigation into the reason why the algorithm produced a bound greater than the maximum should help fix any problems. One drawback in the testing was that the program was only checked against a few tight Bell Inequalities, it would be useful to check the algorithm against Bell Inequalities that are not tight to verify that it works correctly in general.

Scenario	Expected		Outcome		
(n,m,d)	Dimension	Bound	Dimension	Bound	Result
(2,2,2)	8	2	8	2	Pass
(2,3,2)	15	4	15	4	Pass
(3,2,2)	26	6	26	6	Pass
(3,3,2)	63	8	63	8	Pass
(4,2,2)	80	9	1	24	Fail
(4,2,2)	80	10	70	10	Fail

Table 2: The Inequalities the program was tested against and the results. The Algorithm passes all of the tests for Inequalities with $n \leq 3$ but fails for Inequalities with more parties.

The timing of the algorithm was also tested for the Inequalities that it did succeed in producing the correct results for. It is thought that the time taken for the algorithm to calculate the dimension of a general Bell Inequality from a scenario (n, m, d) would take the form (19).

$$\Delta t \propto d^{nm} \cdot k\left((m+1)^n - 1\right) \cdot \left(\frac{m}{m+1}\right)^2 d \tag{19}$$

The reasoning behind this is as follows. The first term d^{nm} is the total number of states the algorithm has to explore, the second term $k((m+1)^n - 1)$ is the total number of the correlators that the algorithm has to calculate for the Inequality, since this varies from Inequality to Inequality the factor $k \in [0, 1]$ will vary from Inequality to Inequality correspondingly. The final term $\left(\frac{m}{m+1}\right)^2 d$ represents the average time taken to calculate each correlator. For each correlator (14) we have to sum over all of the possible values of the outcomes d_i for each party that makes measurements. Suppose there are $\tilde{n}$ parties that make measurements, since there are a total of d outcomes for each of these parties the number of correlator terms that have to be calculated is $\tilde{n}d$, but the size of each correlator term also scales with the number of parties that make measurements. Overall, the time taken to calculate all of the correlator terms in (14) should scale with $\tilde{n}^2 d$. However, the number of parties that make measurements depends on the exact form of each Inequality and so instead of using this, we can replace it with the average number of parties that make measurements. Suppose we consider a random correlator term from a random Inequality, we can ask ourselves what is the average number of parties that make measurements. Since each correlator term will be equally likely and there are m times when each party can make a measurement compared to 1 time when they do not, this will be $m/(m+1)$, we can replace $\tilde{n}$ with this as a rough estimate.

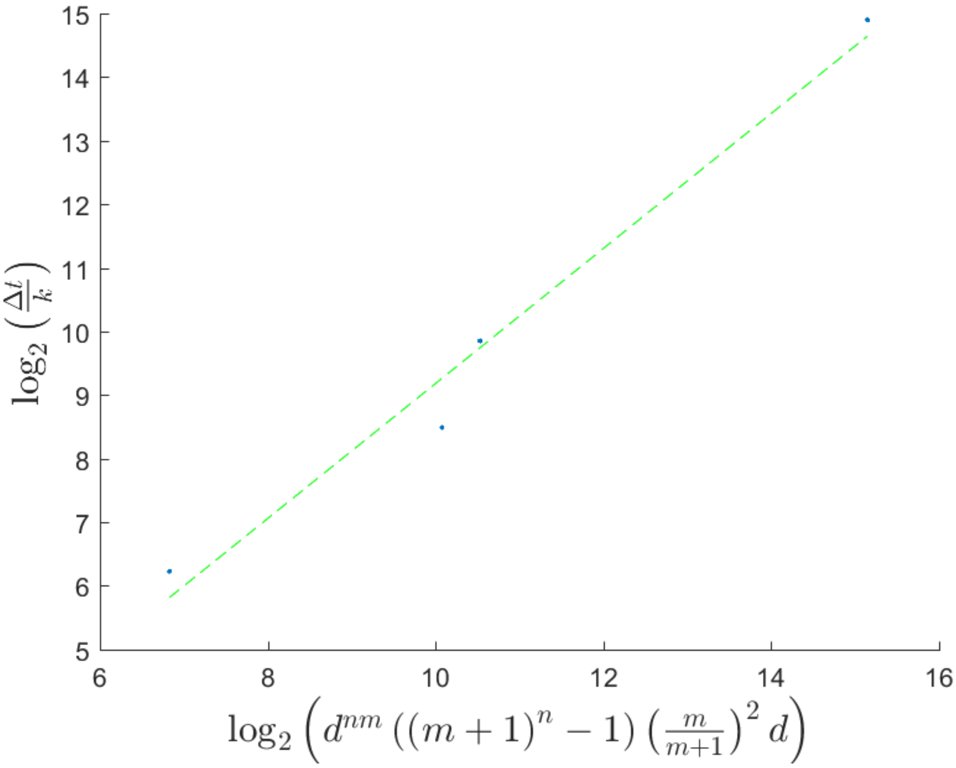

Figure 6: The measured time taken by the algorithm against the expected timings for the tests it passed, both expressed as logarithms, fitted with a straight line. The measured time is the median time taken normalised with respect to k, to account for different Bell Inequalities having different numbers of correlators to calculate.

The algorithm was run several times using the *timeit* function and the median time taken in ms was found. Then, this time was normalised by dividing by the fraction k of coefficients for that Inequality. Since the power dependence on the expected time may also be different, the logarithms of this normalised time and expected time were plotted against each other as shown in Figure 6. A straight line has been fitted to the data using *polyfit*, a least squares algorithm. The gradient tells us the true power dependence and the coefficient of proportionality in (19) can be calculated from the vertical offset. Equation (20) shows the terms plotted, A is the gradient and B is the offset.

$$\log_2(\frac{\Delta t}{k}) = A\log_2\left(d^{nm}\cdot((m+1)^n-1)\cdot\left(\frac{m}{m+1}\right)^2 d\right) + B \tag{20}$$

Exponentiating both sides of (20) leads to (21).

$$\frac{\Delta t}{k} = 2^B\left[d^{nm}\cdot((m+1)^n-1)\cdot\left(\frac{m}{m+1}\right)^2 d\right]^A \tag{21}$$

substituting the results $A = 1.06$ and $B = -1.43$ into (21) we end up with (22), which shows the approximate time scaling of the algorithm with the parameters of the scenario.

$$\frac{\Delta t}{k} \approx 0.37 \left[d^{nm} \cdot ((m+1)^n - 1) \cdot \left(\frac{m}{m+1} \right)^2 d \right]^{1.06} \tag{22}$$

There are still some ways in which the algorithm could be optimized, one idea is to modify the code to skip the calculation of terms with small coefficients as these are negligible. The algorithm's functionality could also be developed further, generalising it to work for any Bell Inequality of any scenario where the parties can make different numbers of measurements and have different numbers of outcomes for each of these measurements.

5 Conclusions

The algorithm works correctly for the simplest scenarios giving both the correct dimension and classical bound, but fails when faced with Inequalities of scenarios with four parties and so cannot be assumed to work properly for any scenario with parties more than this. For one of the Inequalities, it fails to predict both the dimension and bound, but surprisingly, it predicts a bound greater than the known maximum. This suggests that there is some problem in the calculation of the bound itself. On the other hand, it correctly predicts the bound for another Inequality but not its dimension. A detailed investigation into why the algorithm gave a bound higher than the maximum should help reveal any problems.

The time taken for the algorithm to calculate the dimension was measured for the various Bell Inequalities that passed the tests. The measured times were compared to the expected times, the results show us that this model of the algorithm timings fits the data quite well.

Ignoring small coefficient terms would be useful in speeding up the algorithm. The algorithm would be greatly improved if it could be generalised to work for any Inequality of any scenario, where the parties can each perform different numbers of measurements and observe different numbers of outcomes for each of these measurements. It would also be useful to check against Inequalities that are not tight, as all of the inequalities used here were. Overall, the algorithm appears very promising but more work is needed to ensure it functions correctly and within a reasonable amount of time.

A Appendix

The no-signalling conditions require that the marginals of the joint probability distribution be well defined. It can be shown that the quantum behaviours satisfy the no-signalling conditions as follows.

$$\begin{aligned}
\sum_{d_n} P(d_1 d_2 ... d_n | m_1 m_2 ... m_n) &= \mathrm{Tr}\left(M_{m_1}^{d_1} \otimes M_{m_2}^{d_2} \otimes ... \otimes \left(\sum_{d_n} M_{m_n}^{d_n} \right) \rho \right) \\
&= \mathrm{Tr}\left(M_{m_1}^{d_1} \otimes M_{m_2}^{d_2} \otimes ... \otimes \mathbb{1}\, \rho \right) \\
&= \mathrm{Tr}\left(M_{m_1}^{d_1} \otimes M_{m_2}^{d_2} \otimes ... \otimes M_{m_{n-1}}^{d_{n-1}}\, \rho_{1,..,n-1} \right) \\
&= P(d_1 d_2 ... d_{n-1} | m_1 m_2 ... m_{n-1})
\end{aligned} \tag{A.1}$$

Where $\rho_{1,...,n-1}$ is the reduced state found by tracing out over the nth party. This was done by marginalizing over the nth party, but a similar derivation can be done for any party or set of parties.

We can combine the fact that the joint probability distributions of the extremal behaviours can be expressed as a product of the single party local deterministic probabilities with no the no-signalling conditions to end up with a simpler constraint to check.

$$\begin{aligned} \forall\, m_k, m'_k \in M_K, \quad \forall\, k \in \{1, ..., n\}, \quad \forall\, d_i, m_i \in D_i, M_i\ (i \neq k) \\ \sum_{d_k} P(d_1...d_k...d_n|m_1...m_k...m_n) = \sum_{d_k} P(d_1...d_k...d_n|m_1...m'_k...m_n) \\ \sum_{d_k} D_1(d_1|m_1)...D_k(d_k|m_k)...D(d_n|m_n) = \sum_{d_k} D_1(d_1|m_1)...D_k(d_k|m'_k)...D(d_n|m_n) \\ \sum_{d_k} D_k(d_k|m_k) = \sum_{d_k} D_k(d_k|m'_k) \quad \forall\, m_k, m'_k \in M_k \quad \forall\, k \in \{1, ..., n\} \end{aligned} \tag{A.2a}$$

We can also combine the fact that the joint probability distributions can be expressed as products with the normalisation conditions to end up with a second constraint.

$$\begin{aligned} \sum_{d_1,...,d_n} P(d_1...d_n|m_1...m_n) = 1 \quad \forall\, m_1, ..., m_n \\ \sum_{d_1}\sum_{d_2}...\sum_{d_n} D_1(d_1|m_1)D_2(d_2|m_2)...D(d_n|m_n) = 1 \quad \forall\, m_1, ..., m_n \\ \sum_{d_1} D_1(d_1|m_1)\cdot\sum_{d_2} D_2(d_2|m_2)\cdot ... \cdot\sum_{d_n} D_n(d_n|m_n) = 1 \quad \forall\, m_1, ..., m_n \end{aligned} \tag{A.2b}$$

The products of the sums over each individual party's local deterministic probabilities must be unity. However, if we study this more closely we can reduce this further. Consider the values that each of these sums could take, we know that each of these sums can in principle take any integer value from 0 to d, the number of items summed over, since each deterministic probability must be either zero or one. However, suppose one of them does take a value greater than 1, say 2, then in order for the overall expression to be unity the rest of the products must be a fraction, here $1/2$. But this is impossible since as we just mentioned each sum can only take integer values. This implies that none of the sums can take a value greater than 1. Also, if one of the sums takes a value 0 then the whole expression would be invalid, these together imply that each of the sums must individually be unity.

$$\sum_{d_k} D_k(d_k|m_k) = 1 \quad \forall\, m_k \in M_k \quad \forall\, k \in \{1, ..., n\} \tag{A.2c}$$

Comparing this to the previous constraint found (A.2a) we can see that this second one (A.2c) is more restrictive than the first and so the first can be discarded. Therefore, in order to check that the behaviours satisfy both the normalisation and no-signalling conditions and so represent real physical extremal behaviours, one only needs to check they obey (A.2c).

Acknowledgements

I would like to thank Pieter and Mark for their support and giving up their time to supervise me during this work. You've been brilliant to work with and I've very much enjoyed working on this research project with you. I want to thank you for all the help and advice you've given me during this time and the hours you've spent teaching me all the skills and knowledge I would need; I really enjoyed the challenge and it's been a very rewarding experience for me as a researcher. I would also like to thank SURE for giving me this opportunity and funding my research, I've really enjoyed taking part in this event and I hope to do more like this in the future.

References

[1] Alain Aspect, Jean Dalibard, and Gérard Roger. Experimental test of bell's inequalities using time-varying analyzers. *Phys. Rev. Lett.*, 49:1804–1807, 1982.

[2] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964.

[3] J.S. Bell. *Speakable and unspeakable in quantum mechanics*. Cambridge University Press, 1987.

[4] Časlav Brukner, Marek Żukowski, Jian-Wei Pan, and Anton Zeilinger. Bell's inequalities and quantum communication complexity. *Phys. Rev. Lett.*, 92:127901, 2004.

[5] Jing-Ling Chen and Dong-Ling Deng. Bell inequality for qubits based on the cauchy-schwarz inequality. *Phys. Rev. A*, 79:012115, 2009.

[6] B. S. Cirel'son. Quantum generalizations of bell's inequality. *Letters in Mathematical Physics*, 4(2):93–100, 1980.

[7] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, 1969.

[8] Daniel Collins and Nicolas Gisin. A relevant two qubit bell inequality inequivalent to the chsh inequality. *Journal of Physics A: Mathematical and General*, 37(5):1775, 2004.

[9] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, 1935.

[10] Artur K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67:661–663, 1991.

[11] Joshua Geller and Marco Piani. Quantifying non-classical and beyond-quantum correlations in the unified operator formalism. *Journal of Physics A: Mathematical and Theoretical*, 47(42):424030, 2014.

[12] Miguel Navascués, Stefano Pironio, and Antonio Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics*, 10(7):073013, 2008.

[13] S. Pironio. *Aspects of quantum non-locality*. PhD thesis, Université Libre de Bruxelles, 2004.

[14] Stefano Pironio. Lifting bell inequalities. *Journal of Mathematical Physics*, 46(6):062112, 2005.

[15] Sandu Popescu and Daniel Rohrlich. Quantum nonlocality as an axiom. *Foundations of Physics*, 24(3):379–385, 1994.

[16] Reinhard F. Werner and Michael M. Wolf. Bell inequalities and entanglement. *Quantum Info. Comput.*, 1(3):1–25, 2001.

[17] Chunfeng Wu, Jing-Ling Chen, L. C. Kwek, and C. H. Oh. Correlation-function bell inequality with improved visibility for three qubits. *Phys. Rev. A*, 77:062309, 2008.